1

ATTORNEY DOCKET NO. 14114.0344U2

SEQUENCE LISTING

<110> The Government of the United States of America, as represented by
the Secretary, Department of Health and Human Services,
c/o Centers for Disease Control and Prevention
Khudyakov, Yury
Fields, Howard

<120> MOSAIC PROTEIN AND RESTRICTION ENDONUCLEASE ASSISTED
LIGATION METHOD FOR MAKING THE SAME

<130> 14114.0344U2

<140> 09/491,146
<141> 01-25-00

<150> 08/921,887
<151> 25-08-97

<160> 55

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 55

<212> DNA

<213> Hepatitis C virus

<400> 1
ccccgaattc aaccgaaacc gcaacgtaaa accaaacgta acaccattcg tcgtc   55

<210> 2
<211> 69
<212> DNA
<213> Hepatitis C virus

<400> 2
ccccggatcc tatttcggac caacgatctg accaccaccc gggaatttaa cgtcctgcgg   60
acgacgaat   69

<210> 3
<211> 54
<212> DNA
<213> Hepatitis C virus

<400> 3
ccccgaattc aaccgaaacc gcaacgtcag accaaacgta acaccaaccg tcgt   54

<210> 4
<211> 70
<212> DNA
<213> Hepatitis C virus

<400> 4
ccccggatcc tatttcggac caacgatctg accaccaccc gggaatttaa cgtcctgcgg   60
acgacggttg   70

<210> 5
<211> 51
<212> DNA
<213> Hepatitis C virus

<400> 5
ccccgaattc aaccgaaacc gcaacgtaaa accaaacgta acacctaccg t   51

<210> 6
<211> 73
<212> DNA
<213> Hepatitis C virus

<400> 6
ccc009gatcc tatttcggac caacgatctg accaccaccc gggaatttaa cgtcctgcgg   60
acgacggtag gtg   73

<210> 7
<211> 55
<212> DNA
<213> Hepatitis C virus

<400> 7
ccccgaattc aaccgaaacc gcaacgtaaa ccgaaccgta acaccaaccg tcgtc   55

<210> 8
<211> 69
<212> DNA
<213> Hepatitis C virus

<400> 8
ccc008gatcc tatttcggac caacgatctg accaccaccc gggaatttaa cgtcctgcgg   60
acgacggtt   69

<210> 9
<211> 63
<212> DNA
<213> Hepatitis C virus

<400> 9
ccccgaattc aaccgaaacc gcaacgtcag ccgaaacgta acaccccgcg tcgtccgcag   60
gac   63

<210> 10
<211> 60
<212> DNA
<213> Hepatitis C virus

<400> 10
ccccggatcc tatttcggac caacgatctg accaccaccc gggaatttaa cgtcctgcgg   60

<210> 11
<211> 55
<212> DNA
<213> Hepatitis C virus

<400> 11
ccccgaattc aaccgaaacc gcaacgtaaa accaaacgta acgctcaccg tcgtc   55

<210> 12
<211> 69
<212> DNA
<213> Hepatitis C virus

<400> 12
ccccggatcc tatttcggac caacgatctg accaccaccc gggaatttaa cgtcctgcgg   60
acgacggtg   69

<210> 13
<211> 55
<212> DNA
<213> Hepatitis C virus

<400> 13
ccccgaattc aaccgaaacc gcaacgtaaa aaccagcgta acaccaaccg tcgtc   55

<210> 14
<211> 69
<212> DNA
<213> Hepatitis C virus

<400> 14
ccccggatcc tatttcggac caacgatctg accaccaccc gggaatttaa cgtcctgcgg   60
acgacggtt   69

<210> 15
<211> 55
<212> DNA
<213> Hepatitis C virus

<400> 15
ccccgaattc aaccgaaacc gcaacgtaaa accaaacgta acaccattcg tcgtc   55

<210> 16
<211> 69
<212> DNA
<213> Hepatitis C virus

<400> 16

```
ccccggatcc tatttcggaa cgtagataac accaccaccc gggaatttaa cgtcctgcgg          60
acgacgaat                                                                  69
```

```
<210> 17
<211> 56
<212> DNA
<213> Hepatitis C virus
```

```
<400> 17
ccccgaattc aaccgaaacc gcaacgtaaa accgaacgta acaccaaccg tcgtcc              56
```

```
<210> 18
<211> 65
<212> DNA
<213> Hepatitis C virus
```

```
<400> 18
ccccggatcc tatttcggac caacgatctg accaccacca gagaaacgaa cgtccggacg          60
acggt                                                                      65
```

```
<210> 19
<211> 54
<212> DNA
<213> Hepatitis C virus
```

```
<400> 19
ccccgaattc aaccgaaacc gaaacgtcag accaaacgta acaccctgcg tcgt                54
```

```
<210> 20
<211> 73
<212> DNA
<213> Hepatitis C virus
```

```
<400> 20
ccccggatcc tatttcggac caacgatctg accaccagcc gggaatttaa cgtttttcgg          60
acgacgacgc agg                                                             73
```

```
<210> 21
<211> 55
<212> DNA
<213> Hepatitis C virus
```

```
<400> 21
ccccgaattc aaccgaaacc gcaacgtaaa accaaacgta aagctcaccg tcgtc               55
```

```
<210> 22
<211> 69
<212> DNA
<213> Hepatitis C virus
```

```
<400> 22
ccccggatcc tatttcggac caacgatctg accaccaccc gggaatttaa cgtcctgcgg          60
```

acgacggtg   69

<210> 23
<211> 28
<212> PRT
<213> Hepatitis C virus

<400> 23
Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Ile Arg Arg Pro Gln
1   5   10   15
Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20   25
<210> 24
<211> 28
<212> PRT
<213> Hepatitis C virus

<400> 24
Pro Lys Pro Gln Arg Gln Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln
1   5   10   15
Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20   25

<210> 25
<211> 28
<212> PRT
<213> Hepatitis C virus

<400> 25
Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Tyr Arg Arg Pro Gln
1   5   10   15
Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20   25
<210> 26
<211> 28
<212> PRT
<213> Hepatitis C virus

<400> 26
Pro Lys Pro Gln Arg Lys Pro Asn Arg Asn Thr Asn Arg Arg Pro Gln
1   5   10   15
Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20   25
<210> 27
<211> 28
<212> PRT
<213> Hepatitis C virus

<400> 27
Pro Lys Pro Gln Arg Gln Pro Lys Arg Asn Thr Pro Arg Arg Pro Gln
1   5   10   15
Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20   25

```
<210> 28
<211> 28
<212> PRT
<213> Hepatitis C virus

<400> 28
Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Ala His Arg Arg Pro Gln
 1               5                  10                  15
Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20                  25
<210> 29
<211> 28
<212> PRT
<213> Hepatitis C virus

<400> 29
Pro Lys Pro Gln Lys Arg Asn Gln Arg Asn Thr Asn Arg Arg Pro Gln
 1               5                  10                  15
Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20                  25
<210> 30
<211> 28
<212> PRT
<213> Hepatitis C virus

<400> 30
Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Ile Arg Arg Pro Gln
 1               5                  10                  15
Asp Val Lys Phe Pro Gly Gly Gly Val Ile Tyr Val
            20                  25
<210> 31
<211> 28
<212> PRT
<213> Hepatitis C virus

<400> 31
Pro Lys Pro Gln Arg Lys Thr Glu Arg Asn Thr Asn Arg Arg Pro Gln
 1               5                  10                  15
Asp Val Arg Phe Ser Gly Gly Gly Gln Ile Val Gly
            20                  25
<210> 32
<211> 28
<212> PRT
<213> Hepatitis C virus

<400> 32
Pro Lys Pro Lys Arg Gln Thr Lys Arg Asn Thr Leu Arg Arg Pro Lys
 1               5                  10                  15
Asn Val Lys Phe Pro Ala Gly Gly Gln Ile Val Gly
            20                  25
<210> 33
<211> 28
```

```
<212> PRT
<213> Hepatitis C virus

<400> 33
Pro Lys Pro Gln Arg Lys Thr Lys Arg Lys Ala His Arg Arg Pro Gln
 1               5                  10                  15
Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20                  25
<210> 34
<211> 82
<212> DNA
<213> Hepatitis C virus


<400> 34
ccccgaattc aagccgccca cataccatac ctagaacaag gaatgcatct cgcagaacaa      60
ttcaaacaaa aggcacttcg tc                                             82


<210> 35
<211> 84
<212> DNA
<213> Hepatitis C virus


<400> 35
ccccggatcc taactagcct cttccatctc atcaaactcc tgatacaaaa cctccctatc      60
cgggataaca gccggacgaa gtgc                                           84


<210> 36
<211> 85
<212> DNA
<213> Hepatitis C virus


<400> 36
ccccgaattc aagctagtca cttaccgtat atcgagcagg gaatgcagtt agctgaacag      60
tttaaacaga aggctctggc ttttg                                          85


<210> 37
<211> 81
<212> DNA
<213> Hepatitis C virus


<400> 37
ccccggatcc taaggccgag cgtcagactc aggaacataa tgagtaggag aaacatgatt      60
accccgagaa gcaaaagcca g                                              81


<210> 38
<211> 81
<212> DNA
<213> Hepatitis C virus


<400> 38
ccccgaattc aacggcctgc gataataccg gatagggagg ttcttcatag ggagtttgac      60
gagatggagg aggcttttgc g                                              81


<210> 39
```

ATTORNEY DOCKET NO. 14114.0344U2

```
<211> 79
<212> DNA
<213> Hepatitis C virus

<400> 39
ccccggatcc tactgcgaag catcagactc aggaacataa tgagccggac taacatgatt      60
cccacgagac gcaaaagcc                                                    79


<210> 40
<211> 83
<212> DNA
<213> Hepatitis C virus

<400> 40
ccccgaattc aatcgcaggc ggcgccttat attgagcagg ctcaggttat tgctcatcag      60
tttaaggaga aggttcttgc ttt                                               83



<210> 41
<211> 83
<212> DNA
<213> Hepatitis C virus

<400> 41
ccccggatcc tacggcttcg cgtccgactc aggaacataa tgagtcggag aatcatgatt      60
accacgagaa gcaaaagcaa gaa                                               83


<210> 42
<211> 79
<212> DNA
<213> Hepatitis C virus

<400> 42
ccccgaattc aaaagccggc gataatccct gaccgtgagg ttctgtatcg tgagtttgat      60
gagatggagg agtcacagc                                                    79


<210> 43
<211> 87
<212> DNA
<213> Hepatitis C virus

<400> 43
ccccggatcc taaaacgcca gagccttctg cttaaactgc tcagcaagca tcataccctg      60
ctcaatgtac ggaagatgct gtgactc                                           87


<210> 44
<211> 76
<212> DNA
<213> Hepatitis C virus

<400> 44
ccccgaattc aagcgtttgc ttctcgtggt aatcatgttg ctccgactca ttatgttacg      60
gagtcagatg ctaagc                                                       76
```

```
<210> 45
<211> 84
<212> DNA
<213> Hepatitis C virus

<400> 45
ccccggatcc tagaaagcct cctccatctc atcatactgc tgataaagaa cctccttatc        60
cggaaccaga gccggcttag catc                                               84


<210> 46
<211> 80
<212> DNA
<213> Hepatitis C virus

<400> 46
ccccgaattc aagctttcgc ttctcgtggt aatcatgttg ctcctacgca ttatgttgtt        60
gagtcagatg cttctgcttc                                                    80


<210> 47
<211> 86
<212> DNA
<213> Hepatitis C virus

<400> 47
ccccggatcc tagaaagcca gaaccttctc cttaaactga ccagcaatag cacgcgtctc        60
gtccatatac ggcagagaag cagaag                                             86


<210> 48
<211> 80
<212> DNA
<213> Hepatitis C virus

<400> 48
ccccgaattc aagctttcgc tagtcgtggg aatcatgtgt cgccgcgtca ttatgtgcct        60
gagtctgagc ctcaggttgt                                                    80


<210> 49
<211> 80
<212> DNA
<213> Hepatitis C virus

<400> 49
ccccggatcc taagaagcct cctccatctc atcaaaagcc tcatacagta tctccttatc        60
cggcgtaaca acaacctgag                                                    80


<210> 50
<211> 52
<212> DNA
<213> Hepatitis C virus

<400> 50
ccccgaattc aagcttctaa ggccgcgctg attgaggagg gtcagcgtat gg               52


<210> 51
```

ATTORNEY DOCKET NO. 14114.0344U2

```
<211> 48
<212> DNA
<213> Hepatitis C virus

<400> 51
ccccggatcc tactggatct tagacttcag catctcagcc atacgctg          48

<210> 52
<211> 352
<212> PRT
<213> Hepatitis C virus

<400> 52
Ala Ala His Ile Pro Tyr Leu Glu Gln Gly Met His Leu Ala Glu Gln
 1               5                  10                  15
Phe Lys Gln Lys Ala Leu Arg Pro Ala Val Ile Pro Asp Arg Glu Val
            20                  25                  30
Leu Tyr Gln Glu Phe Asp Glu Met Glu Glu Ala Ser His Leu Pro Tyr
        35                  40                  45
Ile Glu Gln Gly Met Gln Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu
    50                  55                  60

Ala Phe Ala Ser Arg Gln Asn His Val Ser Pro Thr His Tyr Val Pro
65                  70                  75                  80
Glu Ser Asp Ala Arg Pro Ala Ile Ile Pro Asp Arg Glu Val Leu His
                85                  90                  95
Arg Glu Phe Asp Glu Met Glu Glu Ala Phe Ala Ser Arg Gly Asn His
            100                 105                 110
Val Ser Pro Ala His Tyr Val Pro Glu Ser Asp Ala Ser Gln Ala Ala
        115                 120                 125
Pro Tyr Ile Glu Gln Ala Gln Val Ile Ala His Gln Phe Lys Glu Lys
    130                 135                 140
Val Leu Ala Phe Ala Ser Arg Gly Asn His Asp Ser Pro Thr His Tyr
145                 150                 155                 160
Val Pro Glu Ser Asp Ala Lys Pro Ala Ile Ile Pro Asp Arg Glu Val
                165                 170                 175
Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu Ser Gln His Leu Pro Tyr
            180                 185                 190
Ile Glu Gln Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu
            195                 200                 205
Ala Phe Ala Ser Arg Gly Asn His Val Ala Pro Thr His Tyr Val Thr
    210                 215                 220

Glu Ser Asp Ala Lys Pro Ala Leu Val Pro Asp Lys Glu Val Leu Tyr
225                 230                 235                 240
Gln Gln Tyr Asp Glu Met Glu Glu Ala Phe Ala Ser Arg Gly Asn His
            245                 250                 255
Val Ala Pro Thr His Tyr Val Val Glu Ser Asp Ala Ser Ala Ser Leu
            260                 265                 270
Pro Tyr Met Asp Glu Thr Arg Ala Ile Ala Gly Gln Phe Lys Glu Lys
        275                 280                 285
Val Leu Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Arg His Tyr
    290                 295                 300
```

ATTORNEY DOCKET NO. 14114.0344U2

```
Val Pro Glu Ser Glu Pro Gln Val Val Val Thr Pro Asp Lys Glu Ile
305                 310                 315                 320
Leu Tyr Glu Ala Phe Asp Glu Met Glu Glu Ala Ser Lys Ala Ala Leu
                325                 330                 335
Ile Glu Glu Gly Gln Arg Met Ala Glu Met Leu Lys Ser Lys Ile Gln
            340                 345                 350
```

<210> 53
<211> 57
<212> DNA
<213> Hepatitis C virus

<400> 53
ctggttccgc gtggatcccc aggaattccc gggtcgactc gagcggccgc atcgtga   57

<210> 54
<211> 27
<212> DNA
<213> Hepatitis C virus

<400> 54
tcgcagcgaa ttctcgagga tccatcc   27

<210> 55
<211> 55
<212> DNA
<213> Hepatitis C virus

<400> 55
ctggttccgc gtggatcgca gcgaattctc gaggatccat ccggccgcat cgtga   55